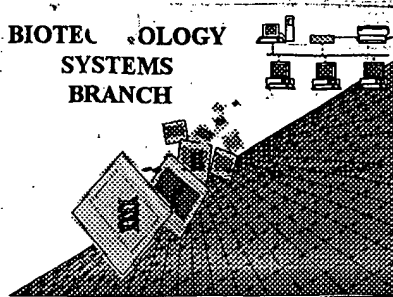


Davis

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



#6

## **RAW SEQUENCE LISTING ERROR REPORT**

BEST AVAILABLE COPY

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/435,274  
Source: 1636  
Date Processed by STIC: 12/28/2000

RECEIVED

JAN 10 2001

TECH CENTER

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin30help@uspto.gov](mailto:patin30help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER  
VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW:

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER:

09/435,274

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4        Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
- 7        PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 8        Skipped Sequences (OLD RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences (NEW RULES)      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10        Use of n's or Xaa's (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11        Use of <213>Organism (NEW RULES)      Sequence(s)        are missing this mandatory field or its response.
- 12        Use of <220>Feature (NEW RULES)      Sequence(s)        are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13        PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

RECEIVED

JAN 10 2001

1636

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/435,274

DATE: 12/28/2000  
TIME: 08:22:34

Input Set : A:\00100301.app  
Output Set : N:\CRF3\12282000\I435274.raw

Does Not Comply  
Corrected Diskette Needed

R2  
RECEIVED

JAN 10 2001

TECH CENTER 1600/2001

3 <110> APPLICANT: Citovsky, Vitaly H  
4 Rhee, Yoon  
6 <120> TITLE OF INVENTION: Genetic Assay for Protein Nuclear Transport  
8 <130> FILE REFERENCE: 001.00301  
10 <140> CURRENT APPLICATION NUMBER: US 09/435,274  
11 <141> CURRENT FILING DATE: 1999-11-05  
13 <150> PRIOR APPLICATION NUMBER: US 60/107,417  
14 <151> PRIOR FILING DATE: 1998-11-06  
16 <160> NUMBER OF SEQ ID NOS: 13  
18 <170> SOFTWARE: PatentIn Ver. 2.1  
20 <210> SEQ ID NO: 1  
21 <211> LENGTH: 611  
22 <212> TYPE: DNA  
23 <213> ORGANISM: Unknown Organism  
25 <220> FEATURE:  
26 <223> OTHER INFORMATION: Description of Unknown Organism:bacterial  
28 <220> FEATURE:  
29 <223> OTHER INFORMATION: modified bacterial lexA  
31 <400> SEQUENCE: 1  
32 atgaaagcgt taacggccag gcaacaagag gtgtttgac tcacccgtga tcacatcagc 60  
33 cagacaggta tgcgccgcag gcctgcgcga atcgcgcagc gtttggggtt cgttccccaa 120  
34 acgcggtctga agaacatctg aaggcgctgg cagcgaagg cgttattgaa attgtttccg 180  
35 gcgcatacag cgggattcgt ctgttgcagg aagaggaaga aggtttgcg ctggtaggtc 240  
36 gtgtgggtgc cgggtgaacca cttctggcgc aacagcatat tgaaggtcat taccaggtcg 300  
37 atccttctct attcaagccg aatgctgatt tcttctgcg cgtcagcggg atgtcgatga 360  
38 aagatatcgg cattatggat ggtgacttgc tggcagtgca taaaactcag gatgtacgta 420  
39 acggtcaggt cgttctgcga cgtattgatg acgaagttac cgttaagggc ctggaaaaac 480  
40 agggcaataa agtcgaactg ttgccagaaa atagcgagtt taaaccaatt gtcgttgacc 540  
41 ttcgtcagca gagcttcacc attgaagggc tggcggttgg ggttatlcgc aacggcgact 600  
42 ggctggaatt c 611  
45 <210> SEQ ID NO: 2  
46 <211> LENGTH: 204  
47 <212> TYPE: PRT  
48 <213> ORGANISM: Unknown Organism  
50 <220> FEATURE:  
51 <223> OTHER INFORMATION: Description of Unknown Organism:bacterial  
53 <220> FEATURE:  
54 <223> OTHER INFORMATION: modified bacterial lexA  
56 <400> SEQUENCE: 2  
57 Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg  
58 1 5 10 15  
60 Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala  
61 20 25 30  
63 Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys  
64 35 40 45  
66 Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg  
67 50 55 60

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/435,274  
 DATE: 12/28/2000  
 TIME: 08:22:34

Input Set : A:\00100301.app  
 Output Set: N:\CRF3\12282000\I435274.raw

```

69 Gly Ile Arg Leu Leu Gln Glu Glu Glu Gly Leu Pro Leu Val Gly
70 65 70 75 80
72 Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly
73 85 90 95
75 His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu
76 100 105 110
78 Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly
79 115 120 125
81 Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val
82 130 135 140
84 Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Gly Leu Glu Lys
85 145 150 155 160
87 Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro
88 165 170 175
90 Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala
91 180 185 190
93 Val Gly Val Ile Arg Asn Gly Asp Trp Leu Glu Phe
94 195 200

```

```

97 <210> SEQ ID NO: 3
98 <211> LENGTH: 7
99 <212> TYPE: PRT
100 <213> ORGANISM: Simian virus 40
102 <220> FEATURE:
103 <223> OTHER INFORMATION: large T antigen NLS
105 <400> SEQUENCE: 3
106 Pro Lys Lys Lys Arg Lys Val
107 1 5
110 <210> SEQ ID NO: 4
111 <211> LENGTH: 16
112 <212> TYPE: PRT
113 <213> ORGANISM: Xenopus sp.
115 <220> FEATURE:
116 <223> OTHER INFORMATION: nucleoplasmin NLS
118 <400> SEQUENCE: 4

```

W--> 119 Lys Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Lys Lys Leu  
 120 1 5 10 15

```

123 <210> SEQ ID NO: 5
124 <211> LENGTH: 9
125 <212> TYPE: PRT
126 <213> ORGANISM: Human immunodeficiency virus type 1
128 <220> FEATURE:
129 <223> OTHER INFORMATION: Rev protein NES
131 <400> SEQUENCE: 5
132 Leu Pro Pro Leu Glu Arg Leu Thr Leu
133 1 5
136 <210> SEQ ID NO: 6
137 <211> LENGTH: 9
138 <212> TYPE: PRT
139 <213> ORGANISM: Artificial Sequence

```

RECEIVED

JAN 10 2001

TECH CENTER 1600/2000

*see Jan 10 on Ena Summary  
Sheet*

RAW SEQUENCE LISTING                      DATE: 12/28/2000  
 PATENT APPLICATION: US/09/435,274        TIME: 08:22:34

Input Set : A:\00100301.app  
 Output Set: N:\CRF3\12282000\I435274.raw

```

141 <220> FEATURE:
142 <223> OTHER INFORMATION: Description of Artificial Sequence:nuclear export
143     signal
145 <220> FEATURE:
146 <223> OTHER INFORMATION: mutated NES of pHEAM10
148 <400> SEQUENCE: 6
149 Leu Pro Pro Asp Leu Arg Leu Thr Leu
150     1                               5
153 <210> SEQ ID NO: 7
154 <211> LENGTH: 4
155 <212> TYPE: PRT
156 <213> ORGANISM: Artificial Sequence
158 <220> FEATURE:
159 <223> OTHER INFORMATION: Description of Artificial Sequence:nuclear export
160     signal
162 <220> FEATURE:
163 <223> OTHER INFORMATION: residual NES of pNEARev(delta)3
165 <400> SEQUENCE: 7
166 Leu Pro Pro Leu
167     1
170 <210> SEQ ID NO: 8
171 <211> LENGTH: 26
172 <212> TYPE: DNA
173 <213> ORGANISM: Artificial Sequence
175 <220> FEATURE:
176 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: GAL4 primer
181 <400> SEQUENCE: 8
182 ggggaattcaa ttttaatcaa agtggg                26
185 <210> SEQ ID NO: 9
186 <211> LENGTH: 27
187 <212> TYPE: DNA
188 <213> ORGANISM: Artificial Sequence
190 <220> FEATURE:
191 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
193 <220> FEATURE:
194 <223> OTHER INFORMATION: GAL4 primer
196 <400> SEQUENCE: 9
197 gacggatccc cgggtattcg atctctt                27
200 <210> SEQ ID NO: 10
201 <211> LENGTH: 29
202 <212> TYPE: DNA
203 <213> ORGANISM: Artificial Sequence
205 <220> FEATURE:
206 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: GAL4 primer
211 <400> SEQUENCE: 10

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/435,274

DATE: 12/28/2000  
TIME: 08:22:34

Input Set : A:\00100301.app  
Output Set: N:\CRF3\12282000\I435274.raw

```
212 gggaattcga taaagcggaa ttaattccc                29
215 <210> SEQ ID NO: 11
216 <211> LENGTH: 27
217 <212> TYPE: DNA
218 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
223 <220> FEATURE:
224 <223> OTHER INFORMATION: GAL4 primer
226 <400> SEQUENCE: 11
227 gacggatccc cgggtattcg atctctt                27
230 <210> SEQ ID NO: 12
231 <211> LENGTH: 25
232 <212> TYPE: DNA
233 <213> ORGANISM: Artificial Sequence
235 <220> FEATURE:
236 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
238 <220> FEATURE:
239 <223> OTHER INFORMATION: mutant lexA primer
241 <400> SEQUENCE: 12
242 ccgttaaggg cctggaaaaa caggg                25
245 <210> SEQ ID NO: 13
246 <211> LENGTH: 26
247 <212> TYPE: DNA
248 <213> ORGANISM: Artificial Sequence
250 <220> FEATURE:
251 <223> OTHER INFORMATION: Description of Artificial Sequence:primer sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: selection lexA primer
256 <400> SEQUENCE: 13
257 gtgactggtg aggcctcaac caagtc                26
```

## VERIFICATION SUMMARY

DATE: 12/28/2000

PATENT APPLICATION: US/09/435,274

TIME: 08:22:35

Input Set : A:\00100301.app

Output Set: N:\CRF3\12282000\I435274.raw

L:119 M:258 W: Mandatory Feature missing, &lt;221&gt; not found for SEQ ID#:4

L:119 M:258 W: Mandatory Feature missing, &lt;222&gt; not found for SEQ ID#:4

L:119 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:4